



## King's Research Portal

DOI:

[10.1155/2017/9586064](https://doi.org/10.1155/2017/9586064)

*Document Version*

Publisher's PDF, also known as Version of record

[Link to publication record in King's Research Portal](#)

*Citation for published version (APA):*

Musmeci, N., Nicosia, V., Aste, T., Di Matteo, T., & Latora, V. (2017). The multiplex dependency structure of financial markets. *COMPLEXITY*. <https://doi.org/10.1155/2017/9586064>

### **Citing this paper**

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

### **General rights**

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

### **Take down policy**

If you believe that this document breaches copyright please contact [librarypure@kcl.ac.uk](mailto:librarypure@kcl.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.

## Research Article

# The Multiplex Dependency Structure of Financial Markets

**Nicolò Musmeci,<sup>1</sup> Vincenzo Nicosia,<sup>2</sup> Tomaso Aste,<sup>3,4</sup>  
Tiziana Di Matteo,<sup>1,3</sup> and Vito Latora<sup>2,5</sup>**

<sup>1</sup>*Department of Mathematics, King's College London, The Strand, London WC2R 2LS, UK*

<sup>2</sup>*School of Mathematical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, UK*

<sup>3</sup>*Department of Computer Science, University College London, Gower Street, London WC1E 6BT, UK*

<sup>4</sup>*Systemic Risk Centre, London School of Economics and Political Sciences, London WC2A 2AE, UK*

<sup>5</sup>*Dipartimento di Fisica ed Astronomia, Università di Catania and INFN, 95123 Catania, Italy*

Correspondence should be addressed to Vito Latora; [v.latora@qmul.ac.uk](mailto:v.latora@qmul.ac.uk)

Received 25 May 2017; Accepted 16 July 2017; Published 20 September 2017

Academic Editor: Tommaso Gili

Copyright © 2017 Nicolò Musmeci et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We propose here a multiplex network approach to investigate simultaneously different types of dependency in complex datasets. In particular, we consider multiplex networks made of four layers corresponding, respectively, to linear, nonlinear, tail, and partial correlations among a set of financial time series. We construct the sparse graph on each layer using a standard network filtering procedure, and we then analyse the structural properties of the obtained multiplex networks. The study of the time evolution of the multiplex constructed from financial data uncovers important changes in intrinsically multiplex properties of the network, and such changes are associated with periods of financial stress. We observe that some features are unique to the multiplex structure and would not be visible otherwise by the separate analysis of the single-layer networks corresponding to each dependency measure.

## 1. Introduction

In the last decade, network theory has been extensively applied to the analysis of financial markets. Financial markets and complex systems in general are comprised of many interacting elements, and understanding their dependency structure and its evolution with time is essential to capture the collective behaviour of these systems, to identify the emergence of critical states, and to mitigate systemic risk arising from the simultaneous movement of several factors. Network filtering is a powerful instrument to associate a sparse network to a high-dimensional dependency measure and the analysis of the structure of such a network can uncover important insights on the collective properties of the underlying system. Following the line first traced by the preliminary work of Mantegna [1], a set of time series associated with financial asset values is mapped into a sparse complex network whose nodes are the assets and whose weighted links represent the dependencies between the corresponding time series. Filtering correlation matrices has been proven to be very useful for the study and characterisation of the underlying

interdependency structure of complex datasets [1–5]. Indeed, sparsity allows filtering out noise, and sparse networks can then be analysed by using standard tools and indicators proposed in complex networks theory to investigate the multivariate properties of the dataset [6, 7]. Further, the filtered network can be used as a sparse inference structure to construct meaningful and computationally efficient predictive models [7, 8].

Complex systems are often characterised by nonlinear forms of dependency between the variables, which are hard to capture with a single measure and are hard to map into a single filtered network. A multiplex network approach, which considers the multilayer structure of a system in a consistent way, is thus a natural and powerful way to take into account simultaneously several distinct kinds of dependency. Dependencies among financial time series can be described by means of different measures, each one having its own advantages and drawbacks, and this has led to the study of different types of networks, namely, correlation networks, causality networks, and so on. The most common approach uses Pearson correlation coefficient to define the weight of a

link, because this is a quantity that can be easily and quickly computed. However, the Pearson coefficient measures the linear correlation between two time series [9], and this is quite a severe limitation, since nonlinearity has been shown to be an important feature of financial markets [10]. Other measures can provide equally informative pictures on assets relationships. For instance, the Kendall correlation coefficient takes into account monotonic nonlinearity [11, 12], while other measures, such as the Tail dependence, quantify dependence in extreme events. It is therefore important to describe quantitatively how these alternative descriptions are related but also differ from the Pearson correlation coefficient and also to monitor how these differences change in time, if at all.

In this work we exploit the power of a multiplex approach to analyse simultaneously different kinds of dependencies among financial time series. The theory of multiplex network is a recently introduced framework that allows describing real-world complex systems consisting of units connected by relationships of different kinds as networks with many layers, where the links at each layer represent a different type of interaction between the same set of nodes [13, 14]. A multiplex network approach, combined with network filtering, is the ideal framework to investigate the interplay between linear, nonlinear, and Tail dependencies, as it is specifically designed to take into account the peculiarity of the patterns of connections at each of the layers but also to describe the intricate relations between the different layers [15].

The idea of analysing multiple layers of interaction was introduced initially in the context of social networks, within the theory of frame analysis [16]. The importance of considering multiple types of human interactions has been more recently demonstrated in different social networks, from terrorist organizations [14] to online communities; in all these cases, multilayer analyses unveil a rich topological structure [17], outperforming single-layer analyses in terms of network modeling and prediction as well [18, 19]. In particular, multilayer community detection in social networks has been shown to be more effective than single-layer approaches [20]; similar results have been reported for community detection on the World Wide Web [21, 22] and citation networks [23]. For instance, in the context of electrical power grids, multilayer analyses have provided important insight into the role of synchronization in triggering cascading failures [24, 25]. Similarly, the analyses on transport networks have highlighted the importance of a multilayer approach to optimize the system against nodes failures, such as flights cancellation [26]. In the context of economic networks, multiplex analyses have been applied to study the World Trade Web [27]. Moreover, they have been extensively used in the context of systemic risk, where graphs are used to model interbank and credit networks [28, 29].

Here, we extend the multiplex approach to financial market time series, with the purpose of analysing the role of different measures of dependencies, namely, the Pearson, Kendall, Tail, and Partial correlation. In particular we consider the so-called Planar Maximally Filtered Graph (PMFG) [2–4, 7] as filtering procedure to each of the four layers. For each of the four unfiltered dependence matrices, the PMFG

filtering starts from the fully connected graph and uses a greedy procedure to obtain a planar graph that connects all the nodes and has the largest sum of weights [3, 4]. The PMFG is able to retain a higher number of links, and therefore a larger amount of information, than the Minimum Spanning Tree (MST) and can be seen as a generalization of the latter which is always contained as a proper subgraph [2]. The topological structures of MST and PMFG have been shown to provide meaningful economic and financial information [30–34] that can be exploited for risk monitoring [35–37] and asset allocation [38, 39]. The advantage of adopting a filtering procedure is not only in the reduction of noise and dimensionality but more importantly in the possibility of generating sparse networks, as sparsity is a requirement for most of the multiplex network measures that will be used in this paper [14]. Other kinds of filtering procedures, including thresholding based methods [35, 40], could have been considered. However, PMFG has the advantage to produce networks with fixed a priori  $(3N - 6)$  number of links that make the comparison between layers and across time windows easier. It is worth mentioning that the filtering of the Partial correlation layer requires an adaptation of the PMFG algorithm to deal with asymmetric relations. We have followed the approach suggested in [41] that rules out double links between nodes. The obtained planar graph corresponding to Partial correlations has been then converted into an undirected graph and included in the multiplex.

## 2. Results

**2.1. Multiplex Network of Financial Stocks.** We have constructed a time-varying multiplex network with  $M = 4$  layers and a varying number of nodes. Nodes represent stocks, selected from a dataset of  $N_{\text{tot}} = 1004$  US stocks which have appeared at least once in S&P500 in the period between 03/01/1993 and 26/02/2015. The period under study has been divided into 200 rolling time windows, each of  $\theta = 1000$  trading days. The network at time  $T = 1, 2, \dots, 200$  can be described by the adjacency matrix  $a_{ij}^\alpha(T)$ , with  $i, j = 1, \dots, N(T)$  and  $\alpha = 1, 2, 3, 4$ . The network at time window  $T$  has  $N(T) < N_{\text{tot}}$  nodes, representing those stocks which were continuously traded in time window  $T$ . The links at each of the four layers are constructed by means of the PMFG procedure from Pearson, Kendall, Tail, and Partial dependencies. The reason for this choice is to provide a complete picture of the market dependency structure: Pearson layer accounts for linear dependency, Kendall layer for monotonic nonlinearity, and Tail dependency for correlation in the tails of returns distribution while Partial correlation detects direct asset-asset relationships which are not explained by the market (see Materials and Methods for details).

Figure 1(a) shows how the average link weight of each of the four dependency networks changes over time. We notice that the average edge weight is a meaningful proxy for the overall level of correlation in one of these dependency layers, since the distribution of edge weights within a layer is normally quite peaked around its mean. The curves shown in Figure 1(a) indicate an overall increase of the typical weights in the examined period 1993–2015 and reveals

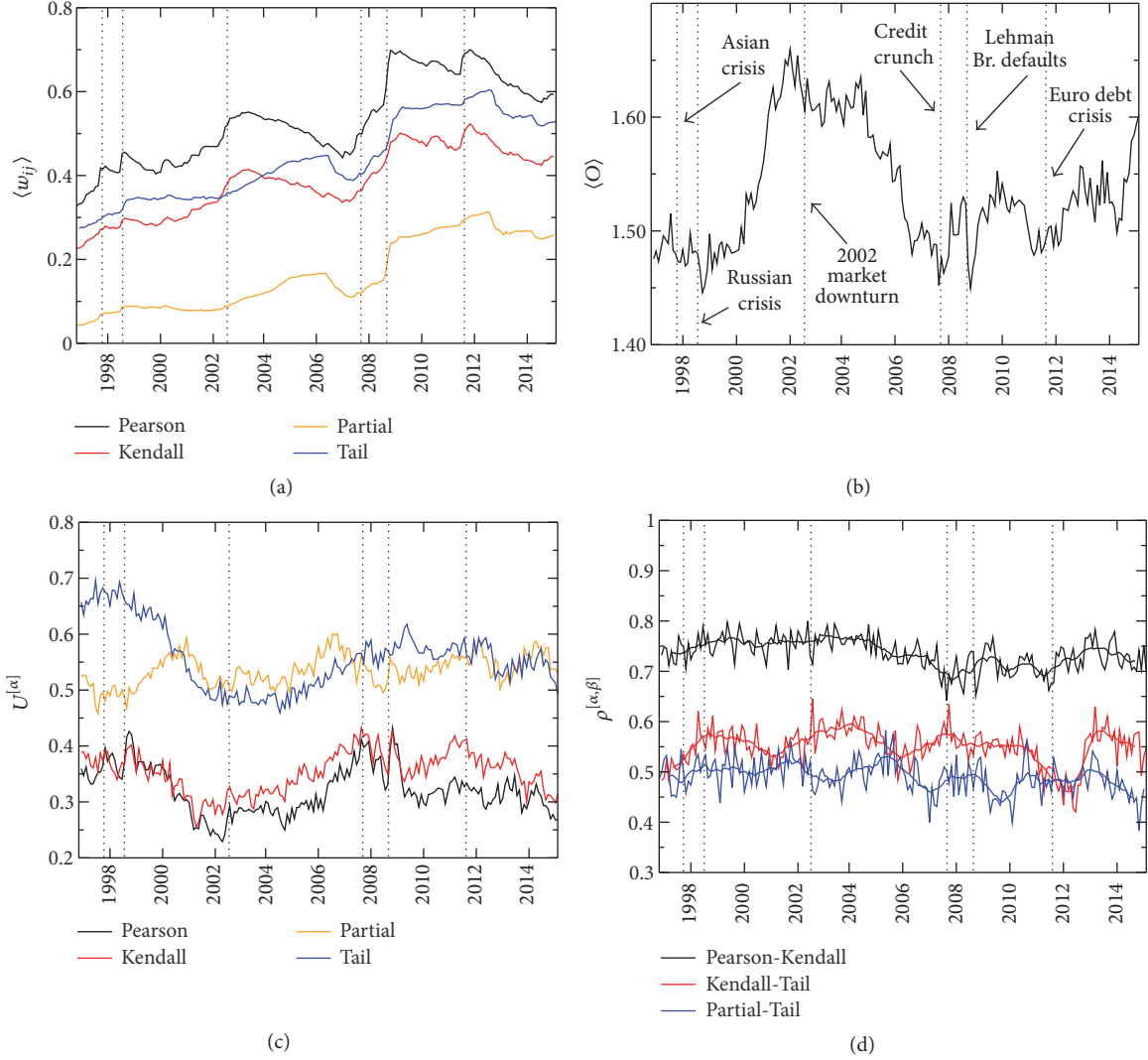


FIGURE 1: The multiplex nature of dependence among financial assets. The plots report the network analysis of a multiplex whose four layers are Planar Maximally Filtered Graphs (PMFGs) obtained from four classical dependence measures, namely, Pearson, Kendall, Tail, and Partial correlation, computed on rolling time windows of 23 trading days between 1993 and 2015. Each of the four layers provides different information on the dependency structure of a market. Although market events and trends have a somehow similar effect on the average dependence  $\langle w_{ij} \rangle$  between nodes at the different layers (panel (a)), each layer has a distinct local structure. This is made evident by the plots of the average edge overlap  $\langle O \rangle$  (panel (b)) and of the fraction  $U^{[\alpha]}$  of edges unique to each layer, which confirm that an edge exists on average on less than two layers, and up to 70% of the edges of a layer are not present on any other layer. Moreover, the same node can have different degrees across the four layers, as indicated by the relatively low values of the pairwise interlayer degree correlation coefficient  $\rho^{[\alpha,\beta]}$  reported in panel (d) for three pairs of layers over the whole observation interval.

a strongly correlated behaviour of the four curves (with linear correlation coefficients between the curves range in  $[0.91, 0.99]$ ). In particular they all display a steep increase in correspondence with the 2007-2008 financial crisis, revealing how the market became more synchronized, regardless of the dependence measure used. This strong correlation in the temporal patterns of the four measures of dependence may lead to the wrong conclusion that the four networks carry very similar information about the structure of financial systems. Conversely, we shall see that even basic multiplex measures suggest otherwise. In Figure 1(b) we report the average edge overlap  $\langle O \rangle$ , that is, the average number of

layers of the financial multiplex network where a generic pair of nodes  $(i, j)$  is connected by an edge (see Materials and Methods for details). Since our multiplex network consists of four layer,  $\langle O \rangle$  takes values in  $[1, 4]$ , and in particular we have  $\langle O \rangle = 1$  when each edge is present only in one layer, while  $\langle O \rangle = 4$  when the four networks are identical. The relatively low values of  $\langle O \rangle$  observed in this case reveal the complementary role played by the different dependency indicators. It is interesting to note that the edge overlap  $\langle O \rangle$  displays a quite dynamic pattern, and its variations seem to be related to the main financial crises highlighted by the vertical lines in Figure 1(b). Overall, what we observe is that periods

of financial turbulence are linked to widening differences among the four layers. Namely, the effect of nonlinearity in the cross-dependence increases, as well as correlation on the tails of returns: the dependence structure becomes richer and more complex during financial crisis. This might be related to the highly nonlinear interactions that characterise investors activities in turbulent periods and that make fat-tail and power-law distributions distinctive features of financial returns. Indeed, if returns were completely described by a multivariate normal distribution, the Pearson layer would be sufficient to quantify entirely the cross-dependence and its relation with the other layers would be trivial and would not change with time. Therefore any variation in the overlapping degree is a signature of increasing complexity in the market. In particular, the first event that triggers a sensible decrease in the average edge overlap is the Russian crisis in 1998, which corresponds to the overall global minimum of  $\langle O \rangle$  in the considered interval. Then,  $\langle O \rangle$  starts increasing towards the end of year 2000 and reaches its global maximum at the beginning of 2002, just before the market downturn of the same year. We observe a marked decrease in 2005, in correspondence with the second phase of the housing bubble, which culminates in the dip associated with the credit crunch at the end of 2007. A second, even steeper drop occurs during the Lehman Brothers default of 2008. After that, the signal appears more stable and weakly increasing, especially towards the end of 2014. Since each edge is present, on average, in less than two layers, each of the four layers effectively provides a partial perspective on the dependency structure of the market. This fact is made more evident by the results reported in Figure 1(c), where we show, for each layer  $\alpha = 1, \dots, 4$ , the fraction of edges  $U^{[\alpha]}$  that exist exclusively in that layer (see Materials and Methods for details). We notice that, at any point in time, from 30% to 70% of the edges of each of the four layers are unique to that layer, meaning that a large fraction of the dependence relations captured by a given measure are not captured by the other measures. For instance, despite the fact that Pearson and Kendall show similar behaviour in Figure 1(c), still between 30% and 40% of the edges on each of those layers exist only on that layer. This indicates that the Pearson and Kendall layers differ for at least 60% to 80% of their edges. In general, each of the four layers is contributing information that cannot be found in the other three layers. It was shown in a recent paper by some of the authors [36] that information filtering networks can be used to forecast volatility outbursts. The present results suggest that a multilayer approach could provide a further forecasting instrument for bear/bull markets. However, this requires further explorations. Interestingly, we observe an increase of  $U^{[\alpha]}$  for all the layers since 2005, which indicates a build-up of nonlinearity and tail correlation in the years preceding the financial crisis: such dynamics might be related to early-risk warnings.

Another remarkable finding is that also the relative importance of a stock in the network, measured for instance by its centrality in terms of degree [39, 42], varies a lot across layers. This is confirmed by the degree correlation coefficient  $\rho^{[\alpha, \beta]}$  for pairs of layers  $\alpha$  and  $\beta$ . In general, high values of

$\rho^{[\alpha, \beta]}$  signal the presence of strong correlations between the degrees of the same node in the two layers (see Materials and Methods for details). Figure 1(d) shows  $\rho^{[\alpha, \beta]}$  as a function of time for three pairs of dependence measures, namely, Pearson-Kendall, Kendall-Tail, and Tail-Partial. Notice that the degrees of the layers corresponding to Pearson and Kendall exhibit a relatively large correlation, which remains quite stable over the whole time interval. Conversely, the degrees of nodes in the Kendall and Tail layers are on average less correlated, and the corresponding values of  $\rho^{[\alpha, \beta]}$  exhibit larger fluctuations. For example, in the tenth time window we find that General Electric stock (GE US) is a hub in Kendall layer with 71 connections, but it has only 16 connections in the Tail layer: therefore the relevance of this stock in the dependence structure depends sensitively on the layer. A similar pattern is observed in the interlayer correlation between the degrees of nodes in Partial and Tail. This might have important implications for portfolio allocation problems, since the asset centrality in the network is related to its risk in the portfolio.

The presence of temporal fluctuations in  $\langle O \rangle$ , in particular the fact that  $\langle O \rangle$  reaches lower values during financial crises, together with the unique patterns of links at each layer, testified by high values of  $U^{[\alpha]}$  and by relatively weak inter-layer degree-degree correlations for some pairs of layers, confirms that an analysis of relations among stocks simply based on one dependence measure can neglect relevant information which can however be captured by other measures. As we will show below, a multiplex network approach which takes into account at the same time all the four dependence measures, but without aggregating them into a single-layer network, is able to provide a richer description of financial markets.

**2.2. Multiedges and Node Multidegrees.** As a first example of useful quantities that can be investigated in a multiplex network, we have computed the so-called multidegree  $k_i^m$  for each node  $i$  in the network, corresponding to different multiedges (see Materials and Methods) [43]. In particular, we have normalised the multidegree of node  $i$  dividing it by the corresponding node overlapping degree  $o_i$ , so that the resulting  $k_i^m/o_i$  is the fraction of multiplex edges of node  $i$  that exist only on a given subset of layers. In Figure 2 we report the average normalised multidegree of each of the 10 industry sectors of the Industry Classification Benchmark (ICB) classification. We focus on the edges existing exclusively in one of the four layers and on the combination of multiedges associated with edges existing in either of the Kendall, Tail, or Partial layer, but not in the Pearson layer. As shown in Figure 2, the multidegree exhibits strong variations in time and high heterogeneity across different industries. Industries such as Oil & Gas, Utilities, and Basic Materials show low values of normalised multidegree in all the four panels (Figures 2(a)–2(d)). Conversely, the edges of nodes corresponding to Industrials, Finance, Technology, Telecommunications, and Consumer Services tend to concentrate in one layer or in a small subset of layers only. For instance, we observe a relatively high concentration of edges at the Kendall layer for nodes corresponding to



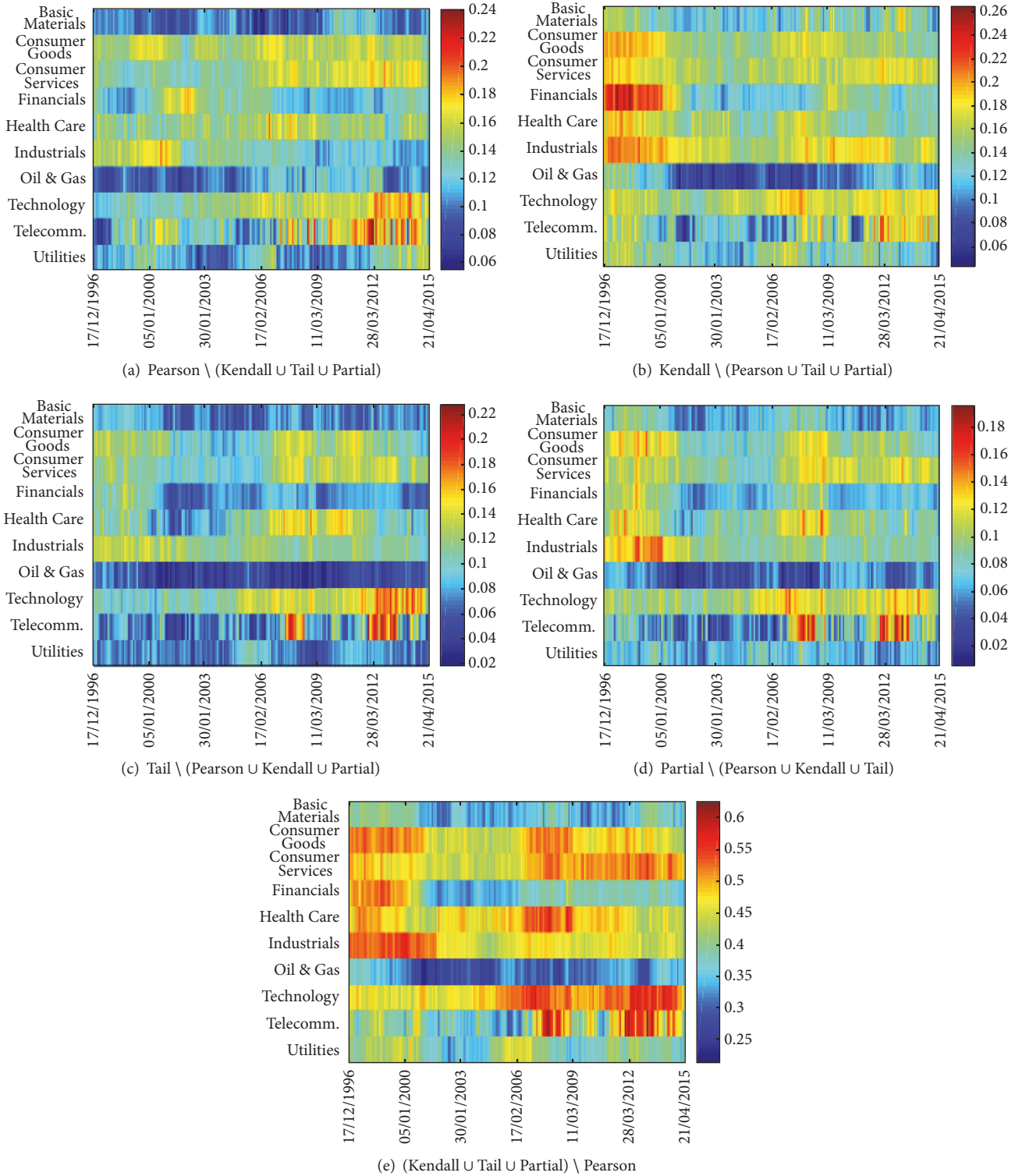


FIGURE 2: Multidegrees reveal the different role of industrial sectors during crises. The plots of the average multidegree of the nodes of the same industrial sector restricted to edges existing exclusively on the (a) Pearson, (b) Kendall, (c) Tail, and (d) Partial layers clearly show that some dependence measures can reveal structures which are unnoticed by other measures. In particular, the plot of the average multidegree associated with edges existing on at least one layer among Kendall, Tail, and Partial, but not on Pearson (panel (e)), reveals that Pearson correlation does not capture many important features such as the prominent role of Basic Materials, Financial, Consumer Goods, and Industrials during crises and the increasing importance of Technology and Consumer Services after the 2007-2008 crisis.

Finance, Industrials, and Consumer Goods stocks in the period preceding the dot-com bubble and the 2002 downturn, a feature not visible in the Pearson layer in Figure 2(a). This implies that for stocks belonging to those industries nonlinearity was a feature of their cross-dependency more important than for other stocks. Analogously, we notice a sudden increase of edges unique to the Tail layer for nodes in Consumer Goods, Consumer Services, and Health Care after the 2007-2008 crisis. Therefore, during this crisis period, the synchronization in the tail region has become a more relevant factor in their dependency structure than before the crisis: this has important implications for portfolio risk, as high Tail dependence can lead to substantial financial losses in case of large price movements. The presence of large heterogeneity and temporal variations in the relative role of different industrial sectors confirms the importance of using a multiplex network approach to analyse dependence among assets. Since industrial sectors have been often used for risk diversification, these findings point out that their use as a diversification benchmark might benefit remarkably from a continuous monitoring based on multiplex: an increase of edges in one layer for an industry can indicate the need of using the corresponding dependency measure for assessing the industry's risk and diversification potential. The fact that different industries display different degrees of nonlinearity and Tail dependence is not surprising after all, given that each industrial sector can be affected in a different way by new information: this industrial specific sensitivity might translate in different cross-dependency properties.

From this perspective it is particularly interesting to discuss the plot of multidegree restricted to edges that are present on either Kendall, Partial, or Tail layer but are not present in the Pearson layer as reported in Figure 2(e). Despite the fact that Pearson correlation coefficient is the most used measure to study dependencies, the plot reveals that until 2002 an analysis of the financial network based exclusively on Pearson correlations would have missed from 40% up to 60% of the edges of assets in sectors such as Basic Materials, Financial, Consumer Goods, and Industrials. The study of evolution with time in Figure 2(e) reveals that the relative role of such industrial sectors in Kendall, Tail, and Partial layers becomes relatively less important between the two crises in 2002 and in 2007, but then such sectors become central again during the 2007-2008 crisis and beyond. This prominent role is quite revealing but it would not had been evident from the analysis of the Pearson layer alone. Let us also note that the period following the 2007-2008 crisis is also characterised by a sensible and unprecedented increase of the normalised multidegree on Kendall, Partial, and Tail layers of stocks belonging to Technology and Telecommunications sectors, whose importance in the market dependence structure has been therefore somehow underestimated over the last ten years by the studies based exclusively on Pearson correlation.

**2.3. Multiplex Cartography of Financial Systems.** To better quantify the relative importance of specific nodes and groups of nodes we computed the overlapping degree and participation coefficient, respectively, measuring the total number of

edges of a node and how such edges are distributed across the layers (see Materials and Methods for details). We started by computing the average degree  $k_I^{[\alpha]}$  at layer  $\alpha$  of nodes belonging to each ICB industry sector  $I$ , defined as  $k_I^{[\alpha]} = (1/N_I) \sum_{i \in I} k_i^{[\alpha]} \delta(c_i, I)$ , where by  $c_i$  we denote the industry of node  $i$  and  $N_I$  is the number of nodes belonging to industry sector  $I$ . Figures 3(a)–3(d) show  $k_I^{[\alpha]}$  as a function of time for each of the four layers.

Notice that nodes in the Financial sector exhibit a quite high average degree, no matter the dependence measure used, with a noticeable peak before the dot-com bubble in 2002. After that, the average degree of Financials has dropped sensibly, with the exception of the 2007-2008 crisis. Apart from the existence of similarities in the overall trend of Financials across the four layers, the analysis of the average degree suggests again the presence of high heterogeneity across sectors and over time.

In the Pearson layer, Basic Materials is the second most central industry throughout most of the observation interval, whereas Industrials and Oil & Gas acquired more connections in the period following the 2007-2008 crisis. The degree in the Kendall layer is distributed more homogeneously among the sectors than in the Pearson layer. Interestingly, the plot of degree on the Tail layer looks similar to that of the Pearson layer. Finally, in the Partial layer we observe the highest level of concentration of links in Finance (consistently to what was found in [41]) and, after the 2007-08 crisis, in Basic Materials.

We have also calculated for each industry  $I$  the average overlapping degree  $o_I \equiv \langle o_i \rangle_{i \in I}$ , where  $o_i$  is the overlapping degree of node  $i$ , which quantifies the overall importance of each industrial sector in the multiplex dependence network. The average overlapping degree of each industry is shown as a function of time in Figure 3(e). As we can see,  $o_I$  is able to highlight the prominent role played in the multiplex network by Financials, Basic Materials, Oil & Gas, and Industrials sectors, revealing also the presence of four different phases between 1997 and 2015. The first phase, during which Financials is the only prominent industry, covers the period between 1997 and 2000. The second phase, between 2000 and the 2007-08 crisis, is characterised by the emergence of Basic Materials as the second most central sector. In the third phase, between 2009 and 2014, Financials loses its importance in favour of Industrials, Oil & Gas, and Basic Materials (that becomes the most central one). Finally, in 2014 a new equilibrium starts to emerge, with Financials and Industrials gaining again a central role in the system.

The participation coefficient complements the information provided by the overlapping degree, quantifying how the edges of a node are distributed over the layers of the multiplex. In particular, the participation coefficient of node  $i$  is equal to 0 if  $i$  has edges in only one of the layers, while it is maximum and equal to 1 when the edges of node  $i$  are equally distributed across the layers (see Materials and Methods for details). In Figure 3(f) we report, as a function of time, the average participation coefficient  $P_I$  for each ICB industry  $I$ . Interestingly, the plot reveals that the increase of the overlapping degrees of Financials, Basic Materials, Industrials,

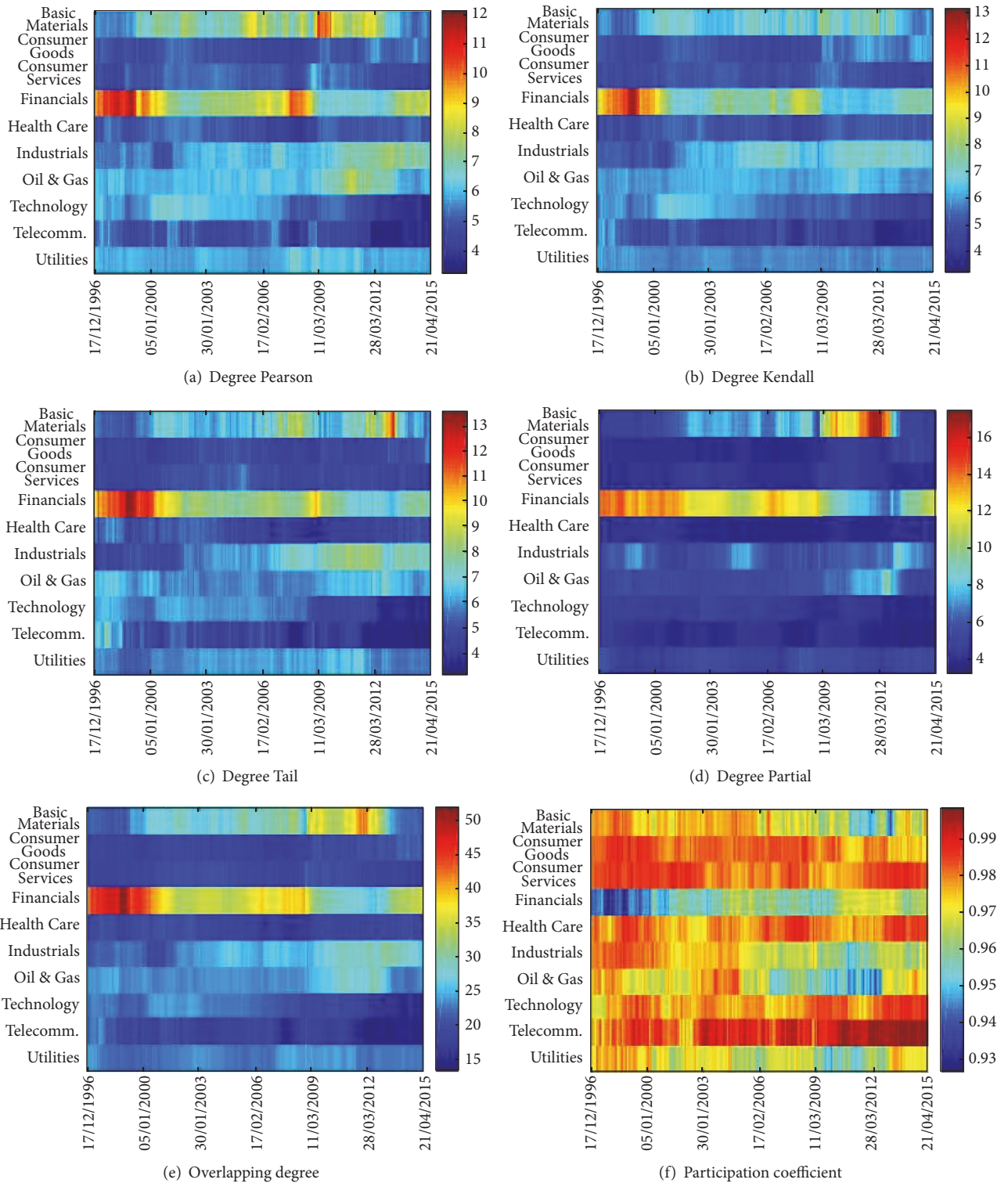


FIGURE 3: Average node degree as a proxy of the importance of an industry. The plots of average degree of the nodes belonging to the different industrial sectors restricted to the (a) Pearson, (b) Kendall, (c) Tail, and (d) Partial layers and of the average overlapping degree reported in panel (e) confirm the relative importance of Financials. However, the average participation coefficient (panel (f)) suggests that the dependence structure of some sectors such as Basic Materials, Industrials, and Oil & Gas has become more heterogeneous, that is, focusing only on a subset of the four layers, after the 2007-2008 crisis.



and Oil & Gas sectors shown in Figure 3(e) is normally accompanied by a substantial decrease of the corresponding participation coefficients. This indicates that those sectors accumulated degree on just one or two layers, confirming what we found in multidegree analysis. A somehow more detailed analysis of the temporal evolution of participation coefficient for each sector is reported in Section 4.5.

### 3. Discussion

By using filtered networks from different correlation measures we have demonstrated that a multiplex network approach can reveal features that would have otherwise been invisible to the analysis of each dependency measure in isolation. Although the layers produced, respectively, from Pearson, Kendall, Tail, and Partial correlations show a certain overall similarity, they exhibit distinct features that are associated with market changes. For instance, we observed that average edge overlap between the first three layers drops significantly during periods of market stress revealing that nonlinear effects are more relevant during crisis periods. The analysis of the average multidegree associated with edges not present on the Pearson layer, but existing on at least one of the three remaining layers, indicates that Pearson correlations alone can miss detecting some important features. We observed that the relative importance of nonlinearity and tails on market dependence structure, as measured by mean edge overlap between the last three layers, has dropped significantly in the first half of the 2000s and then risen steeply between 2005 and the 2007-08 crisis. Overall, financial crises trigger remarkable drops in the edge overlap, widening therefore the differences among the measures of dependence just when evaluation of risk becomes of the highest importance. Different industry sectors exhibit different structural overlaps. For instance, Financials, Industrials, and Consumer Goods show an increasing number of connections only on Kendall layer in the late 90s/early 2000s, at the edge of the dot-com bubble. After the 2007-08 crisis these industries tend to have many edges on the Kendall, Tail, and Partial which are not present on the Pearson layer. This observation questions whether we can rely on the Pearson estimator alone, when analysing correlations between stocks. A study of the overlapping degree and of the participation coefficient shows that asset centrality, an important feature for portfolio optimization [39, 40], changes considerably across layers with largest desynchronized changes occurring during periods of market distress. Overall our analysis indicates that different dependency measures provide complementary pieces of information about the structure and evolution of markets and that a multiplex network approach can be useful in capturing systemic properties that would otherwise go unnoticed.

### 4. Materials and Methods

**4.1. Dataset.** The original dataset consists of the daily prices of  $N_{\text{tot}} = 1004$  US stocks traded in the period between 03/01/1993 and 26/02/2015. Each stock in the dataset has been included in S&P500 at least once in the period considered.

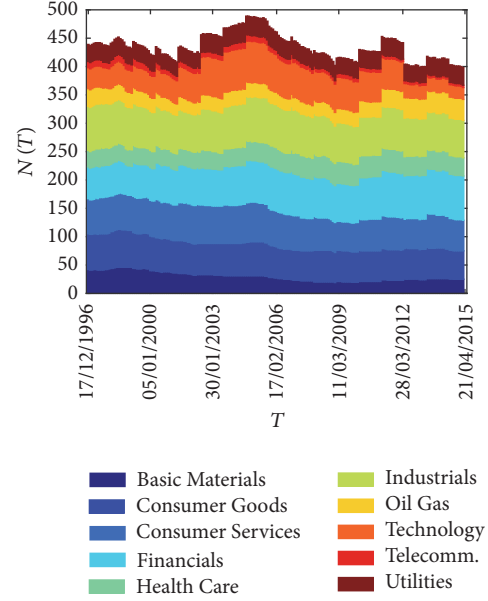


FIGURE 4: Number of stocks in each ICB industry in time. Number of stocks that are continuously traded in each time window together with their partition in terms of ICB industries.

Hence the stocks considered provide a representative picture of the US stock market over an extended time window of 22 years and cover all the 10 industries listed in the Industry Classification Benchmark (ICB) (Figure 4). It is important to notice that most of the stocks in this set are not traded over the entire period. This is a major difference with respect to the majority of the works on dynamic correlation-based networks, in which only stocks continuously traded over the period under study are considered, leading to a significant “survival bias.” For each asset  $i$  we have calculated the series of log-returns, defined as  $r_i(t) = \log(P_i(t)) - \log(P_i(t-1))$ , where  $P_i(t)$  is stock price at day  $t$ . The construction of the time-varying multiplex networks is based on log-returns and has been performed in moving time windows of  $\theta = 1000$  trading days (about 4 years), with a shift of  $dT = 23$  trading days (about one month), adding up to 200 different multiplex networks, one for each time window. For each time window  $T$ , four different  $N(T) \times N(T)$  dependence matrices have been computed, respectively, based on the four different estimators illustrated in Section 4.2. Since the number of active stocks changes with time, dependence matrices at different times can have different number of stocks  $N(T)$ , as shown in Figure 4. In the figure the ICB industry composition of our dataset in each time window is also shown, confirming that we have a representative sample of all market throughout the period. We have verified that the results we are discussing in the following are robust against change of  $\theta$  and  $dT$ .

**4.2. Dependence among Financial Time Series.** We have considered four different measures of dependence between two time series  $r_i(u)$  and  $r_j(u)$ ,  $i, j = 1, 2, \dots, N$ ,  $u = 1, 2, \dots, \theta$ , indicated in the following, respectively, as Pearson, Kendall, Tail, and Partial.

**4.2.1. Pearson Dependence.** It is a measure of linear dependence between two time series and is based on the evaluation of the Pearson correlation coefficient [44]. We have used the exponentially smoothed version of this estimator [45], in order to mitigate excessive sensitiveness to outliers in remote observations:

$$\rho_{ij}^w = \frac{\sum_{u=1}^{\theta} w_u (r_i(u) - \bar{r}_i^w) (r_j(u) - \bar{r}_j^w)}{\sqrt{\sum_{u=1}^{\theta} w_u (r_i(u) - \bar{r}_i^w)^2} \sqrt{\sum_{u=1}^{\theta} w_u (r_j(u) - \bar{r}_j^w)^2}}, \quad (1)$$

with

$$w_u = w_0 \exp\left(\frac{u - \theta}{T^*}\right), \quad (2)$$

where  $T^*$  is the weight characteristic time ( $T^* > 0$ ) that controls the rate at which past observations lose importance in the correlation and  $w_0$  is a constant connected to the normalisation constraint  $\sum_{u=1}^{\theta} w_u = 1$ . We have chosen  $T^* = \theta/3$  according to previously established criteria [45].

**4.2.2. Kendall Dependence.** It is a measure of dependence that takes into account the nonlinearity of a time series. It is based on the evaluation of the so-called Kendall's  $\tau$  rank correlation coefficient, starting from the quantities  $d_k(u, v) \equiv \text{sgn}(r_k(u) - r_k(v))$ . The estimator counts the number of concordant pairs, that is, pairs of observations such that  $d_i(u, v)$  and  $d_j(u, v)$  have equal signs, minus the number of discordant pairs [11]. As for the case of the Pearson dependence, we have used the exponentially smoothed version of the estimator [45]:

$$\tau_{ij}^w = \sum_{u=1}^{\theta} \sum_{v=u+1}^{\theta} w_{u,v} d_i(u, v) d_j(u, v), \quad (3)$$

with

$$w_{u,v} = w_0 \exp\left(\frac{u - \theta}{T^*}\right) \exp\left(\frac{v - \theta}{T^*}\right), \quad (4)$$

where  $T^*$  is again the weight characteristic time.

**4.2.3. Tail Dependence.** It is a nonparametric estimator of tail copula that provides a measure of dependence focused on extreme events. It is based on the evaluation of the following estimator [46]:

$$C_{ij}(p_1, p_2) = \frac{\sum_{u=1}^{\theta} \mathbb{1}_{\{F^i(r_i(u)) < p_1 \wedge F^j(r_j(u)) < p_2\}}}{\sum_{u=1}^{\theta} \mathbb{1}_{\{F^i(r_i(u)) < p_1 \vee F^j(r_j(u)) < p_2\}}}, \quad (5)$$

where  $F^i$  and  $F^j$  are the empirical cumulative probabilities of returns  $r_i(u)$  and  $r_j(u)$ , respectively, and  $p_1$  and  $p_2$  are two parameters representing the percentiles above which an observation is considered (lower) tail. We focus on lower tails since we are interested in risk management applications, where the attention is on losses. It can be shown that this is a consistent estimator of tail copula [46]. In this work we have chosen  $p_1 = p_2 = 0.1$  (i.e., we consider tail every observation below the 10th percentile), as a trade-off between the need of statistic and the interest in extreme events.

**4.2.4. Partial Dependence.** It is a measure of dependence that quantifies to what extent each asset affects other assets correlation. The Partial correlation  $\rho_{ik|j}$ , or correlation influence, between assets  $i$  and  $k$  based on  $j$ , is the Pearson correlation between the residuals of  $r_i(u)$  and  $r_k(u)$  obtained after regression against  $r_j(u)$  [47]. It can be written in terms of a Pearson correlation coefficient as follows [41]:

$$\rho_{ik|j} = \frac{\rho_{ik} - \rho_{ij}\rho_{kj}}{\sqrt{[1 - \rho_{ij}^2][1 - \rho_{kj}^2]}}. \quad (6)$$

This measure represents the amount of correlation between  $i$  and  $k$  that is left once the influence of  $j$  is subtracted. Following [41], we define the correlation influence of  $j$  on the pair  $i, k$  as

$$d(i, k | j) = \rho_{ik} - \rho_{ik|j}. \quad (7)$$

$d(i, k | j)$  is large when a significant fraction of correlation between  $i$  and  $k$  is due to the influence of  $j$ . Finally, in order to translate this into a measure between  $i$  and  $j$ , the so-called Partial dependence, we average it over the index  $k$ :

$$d(i | j) = \langle d(i, k | j) \rangle_{k \neq i, j}. \quad (8)$$

$d(i | j)$  is the measure of influence of  $j$  on  $i$  based on Partial correlation. It is worth noting that, unlike the other measures of dependence,  $d(i | j)$  provides a directed relation between assets (as in general  $d(i | j) \neq d(j | i)$ ). In the rest of the paper we refer to this indicator as ‘‘Partial dependence,’’ even though strictly speaking we are analysing the correlation influence based on Partial correlation.

**4.3. Graph Filtering and the Construction of the Multiplex Network.** For each of the 200 time windows we have then constructed a multiplex network with  $M = 4$  layers obtained, respectively, by means of the four dependence indicators. In order to reduce the noise and the redundancy contained in each dependence matrix we have applied the Planar Maximally Filtered Graph [2–4, 7]. It is worth mentioning that the filtering of the correlation influence layer requires an adaptation of the PMFG algorithm to deal with asymmetric relations. We have followed the approach suggested in [41] that rules out double links between nodes. The obtained planar graphs have been then converted into undirected graphs and included in the multiplex.

**4.4. Multiplex Measures.** Let us consider a weighted multiplex network  $\mathcal{M}$  on  $N$  nodes, defined by the  $M$ -dimensional array of weighted adjacency matrices  $\mathcal{W} = \{W^{[1]}, W^{[2]}, \dots, W^{[M]}\}$ , where  $W^{[\alpha]} = \{w_{ij}^{[\alpha]}\}$  are the matrices of weights that determine the topology of the  $\alpha$ th layer though the PMFG filtering. Here the weight  $w_{ij}^{[\alpha]}$  represents the strength of the correlation between node  $i$  and node  $j$  on layer  $\alpha$ , where the different layers are obtained through different correlation measures. In the following we will indicate by  $W^{[\alpha]}$  the weighted adjacency matrix of the PMFG associated with layer  $\alpha$  and by  $A^{[\alpha]}$  the corresponding unweighted

adjacency matrix, where  $a_{ij}^{[\alpha]} = 1$  if and only if  $w_{ij}^{[\alpha]} \neq 0$ . We denote by  $K^{[\alpha]} = (1/2) \sum_{i,j} a_{ij}^{[\alpha]}$  the number of edges on layer  $\alpha$  and by  $K = (1/2) \sum_{i,j} [1 - \prod_{\alpha} (1 - a_{ij}^{[\alpha]})]$  the number of pairs of nodes which are connected by at least one edge on at least one of the  $M$  layers. Notice that since the network at each layer is a PMFG, then we have  $K^{[\alpha]} = 3(N-2) \forall \alpha$  by construction.

We consider some basic quantities commonly used to characterise multiplex networks [14, 43]. The first one is the mean edge overlap, defined as the average number of layers on which an edge between two randomly chosen nodes  $i$  and  $j$  exists:

$$\langle O \rangle = \frac{1}{2K} \sum_{i,j} \sum_{\alpha} a_{ij}^{[\alpha]}. \quad (9)$$

Notice that  $\langle O \rangle = 1$  only when all the  $M$  layers are identical; that is,  $A^{[\alpha]} \equiv A^{[\beta]} \forall \alpha, \beta = 1, \dots, M$ , while  $\langle O \rangle = 0$  if no edge is present in more than one layer, so that the average edge overlap is in fact a measure of how much similar the structures of the layers of a multiplex network are. A somehow dual quantity is the fraction of edges of layer  $\alpha$  which do not exist on any other layer:

$$U^{[\alpha]} = \frac{1}{2K^{[\alpha]}} \sum_{i,j} a_{ij}^{[\alpha]} \prod_{\beta \neq \alpha} (1 - a_{ij}^{[\beta]}) \quad (10)$$

which quantifies how peculiar the structure of a given layer  $\alpha$  is, since  $U^{[\alpha]}$  is close to zero only when almost all the edges of layer  $\alpha$  are also present on at least one of the other  $M-1$  layers.

More accurate information about the contribution of each node to a layer (or to a group of layers) can be obtained by the so-called multidegree of a node  $i$ . Let us consider the vector  $\vec{m} = (m_1, m_2, \dots, m_M)$ , with  $M$  equal to the number of layers, where each  $m_{\alpha}$  can take only two values  $\{1, 0\}$ . We say that a pair of nodes  $i, j$  has a multilink  $\vec{m}$  if they are connected only on those layers  $\alpha$  for which  $m_{\alpha} = 1$  in  $\vec{m}$  [43]. The information on the  $M$  adjacency matrices  $a_{ij}^{[\alpha]}$  ( $\alpha = 1, \dots, M$ ) can then be aggregated in the multiadjacency matrix  $A_{ij}^{\vec{m}}$ , where  $A_{ij}^{\vec{m}} = 1$  if and only if the pair  $i, j$  is connected by a multilink  $\vec{m}$ . Formally [13, 43]

$$A_{ij}^{\vec{m}} \equiv \prod_{\alpha=1}^M [a_{ij}^{\alpha} m_{\alpha} + (1 - a_{ij}^{\alpha})(1 - m_{\alpha})]. \quad (11)$$

From the multiadjacency matrix we can define the multidegree  $\vec{m}$  of a node  $i$ , as the number of multilinks  $\vec{m}$  connecting  $i$ :

$$k_i^{\vec{m}} = \sum_j A_{ij}^{\vec{m}}. \quad (12)$$

This measure allows us to calculate, for example, how many edges node  $i$  has on layer 1 only ( $k_i^{\vec{m}}$  choosing  $m_1 = 1, m_{\alpha} = 0, \forall \alpha \neq 1$ ), integrating the global information provided by  $U^{[\alpha]}$ .

The most basic measure to quantify the importance of single nodes on each layer is by means of the node degree  $k_i^{[\alpha]} = \sum_j a_{ij}^{[\alpha]}$ . However, since the same node  $i$  is normally present at all layers, we can introduce two quantities to characterise the role of node  $i$  in the multiplex [14], namely, the overlapping degree

$$o_i = \sum_{\alpha} k_i^{[\alpha]} \quad (13)$$

and the multiplex participation coefficient:

$$P_i = \frac{M}{M-1} \left[ 1 - \sum_{\alpha} \left( \frac{k_i^{[\alpha]}}{o_i} \right) \right]. \quad (14)$$

The overlapping degree is just the total number of edges incident on node  $i$  at any layer, so that node are classified as *hubs* if they have a relatively large value of  $o_i$ . The multiplex participation coefficient quantifies the dispersion of the edges incident on node  $i$  across the layers. In fact,  $P_i = 0$  if the edges of  $i$  are concentrated on exactly one of the  $M$  layers (in this case  $i$  is a *focused* node), while  $P_i = 1$  if the edges of  $i$  are uniformly distributed across the  $M$  layers, that is, when  $k_i^{[\alpha]} = o_i/M \forall \alpha$  (in which case  $i$  is a *truly multiplex* node). The scatter plot of  $o_i$  and  $P_i$  is called *multiplex cartography* and has been used as a synthetic graphical representation of the overall heterogeneity of node roles observed in a multiplex.

In a multiplex network, it is important also to look at the presence and sign of interlayer degree correlations. This can be done by computing the interlayer degree correlation coefficient [15]:

$$\rho^{[\alpha, \beta]} = \frac{\sum_i (R_i^{[\alpha]} - \overline{R^{[\alpha]}})(R_i^{[\beta]} - \overline{R^{[\beta]}})}{\sqrt{\sum_i (R_i^{[\alpha]} - \overline{R^{[\alpha]}})^2 \sum_j (R_j^{[\beta]} - \overline{R^{[\beta]}})^2}}, \quad (15)$$

where  $R_i^{[\alpha]}$  is the rank of node  $i$  according to its degree on layer  $\alpha$  and  $\overline{R^{[\alpha]}}$  is the average rank by degree on layer  $\alpha$ . In general  $\rho^{[\alpha, \beta]}$  takes values in  $[-1, 1]$ , where values close to  $+1$  and  $-1$ , respectively, indicate the strong positive and negative correlations, while  $\rho^{[\alpha, \beta]} \approx 0$  if the degrees at the two layers are uncorrelated.

**4.5. Time Evolution of the Average Participation Coefficient.** In Figure 5 we plot the time evolution of the average participation coefficient  $P_I$  ( $x$ -axis) of stocks in the industrial sector  $I$  against the average overlapping degree  $o_I$  ( $y$ -axis). Each circle corresponds to one of the 200 time windows, while the size and colour of each circle represent different time windows. Each panel corresponds to one industrial sector  $I$ . The diagrams reveal that in the last 20 years the role of different sectors has changed radically and in different directions. For instance, stocks in the Financials sector evolved from a relatively large overlapping degree and a small participation coefficient in the late 1990s to a smaller number of edges, distributed more homogeneously across the layers, towards the end of the observation period. Conversely, Industrials

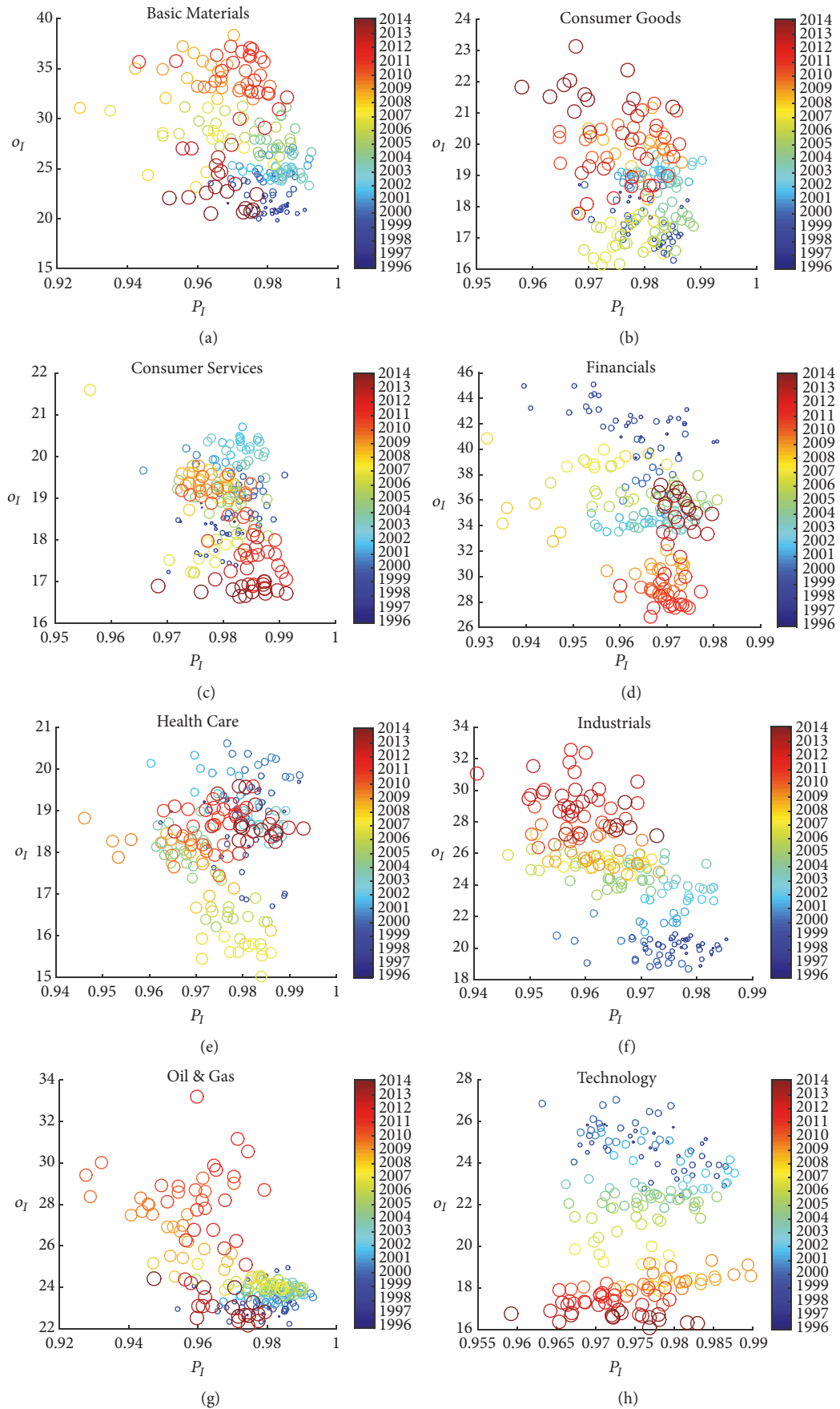


FIGURE 5: Continued.



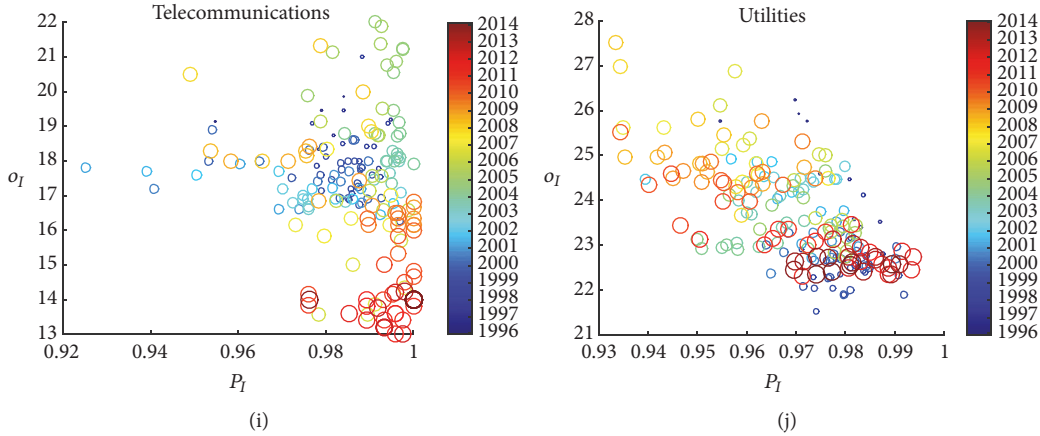


FIGURE 5: Industries evolution in the overlapping degree/participation coefficient plane. Fixing an industry  $I$ , we have plotted for each time window a circle whose  $y$  coordinate is the average overlapping degree  $o_I$  and whose  $x$  coordinate is the average participation coefficient  $P_I$ . Points at different times are characterised by different sizes (small to large) and colours (legend on the right). In (a)–(j) we show the results, respectively, for Basic Materials, Consumer Goods, Consumer Services, Financials, Health Care, Industrials, Oil & Gas, Technology, Telecommunications, and Utilities.

stocks have acquired degree on some of the layers, resulting in a considerable decrease of participation coefficient. This is another indication of the importance of monitoring all the layers together, as an increase in the structural role of an industry (as measured by the overlapping degree) is typically due to only a subset of layers (as indicated by the corresponding decrease of the participation coefficient).

## Conflicts of Interest

The authors declare no competing financial interests.

## Authors' Contributions

Nicolò Musmeci and Vincenzo Nicosia contributed equally to this work. Nicolò Musmeci, Vincenzo Nicosia, Tomaso Aste, Tiziana Di Matteo, and Vito Latora devised the study, performed the experiments and simulations, analysed the results, wrote the paper, and approved the final draft.

## Acknowledgments

The authors wish to thank Alessandro Fiasconaro for useful discussions at the beginning of this project. Vito Latora acknowledges support from the EPSRC Projects EP/K020633/1 and EP/N013492/1. The authors wish to thank Bloomberg for providing the data. Tiziana Di Matteo wishes to thank the COST Action TD1210 for partially supporting this work.

## References

- [1] R. Mantegna, "Hierarchical structure in financial markets," *The European Physical Journal B*, vol. 11, no. 1, pp. 193–197, 1999.
- [2] M. Tumminello, T. Aste, T. Di Matteo, and R. Mantegna, "A tool for filtering information in complex systems," *Proc. Natl. Acad. Sci.*, vol. 102, pp. 10421–10426, 2005.
- [3] T. Aste, T. Di Matteo, and S. T. Hyde, "Complex networks on hyperbolic surfaces," *Physica A* 346, 2005.
- [4] T. Aste and T. Di Matteo, "Dynamical networks from correlations," *Physica A: Statistical Mechanics and its Applications*, vol. 370, no. 1, pp. 156–161, 2006.
- [5] T. Aste, R. Grammatica, and T. Di Matteo, "Exploring complex networks via topological embedding on surfaces," *Phys. Rev. E*, vol. 86, Article ID 036109, 2012.
- [6] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. W. Hwang, "Complex networks: Structure and dynamics," *Physics Reports*, vol. 424, no. 4–5, pp. 175–308, 2006.
- [7] G. P. Massara, T. Di Matteo, and T. Aste, "Network filtering for big data: Triangulated maximally filtered graph," *J Complex Networks*, vol. 5, no. 2, 2017.
- [8] W. Barfuss, G. P. Massara, T. Di Matteo, and T. Aste, "Parsimonious modeling with information filtering networks," *Covering Statistical, Nonlinear, Biological, and Soft Matter Physics*, vol. 94, no. 6, 2016.
- [9] P. Embrechts, A. McNeil, and D. Straumann, "Correlation and dependency in risk management: properties and pitfalls," in *Risk Management: Value at Risk and Beyond*, M. Dempster et al., Ed., Cambridge University Press, Cambridge, UK, 2001.
- [10] D. Sornette and J. Andersen, "A nonlinear super-exponential rational model of speculative financial bubbles," *Int. J. Mod. Phys. C*, vol. 13, pp. 171–188, 2002.
- [11] M. G. Kendall, "A New Measure of Rank Correlation," *Biometrika*, vol. 30, no. 1/2, p. 81, 1938.
- [12] G. Meissner, *Correlation Risk Modeling and Management*, Harvard University Press, 2014.
- [13] S. Boccaletti et al., "The structure and dynamics of multilayer networks," *Phys Rep*, vol. 544, pp. 1–122, 2014.
- [14] F. Battiston, V. Nicosia, and V. Latora, "Metrics for the analysis of multiplex networks," *Phys. Rev. E*, vol. 89, Article ID 032804, 2013.
- [15] V. Nicosia and V. Latora, "Measuring and modelling correlations in multiplex networks," *Phys. Rev. E*, vol. 92, Article ID 032805, 2015.

- [16] E. Goffman, *Frame Analysis: An Essay on the Organization of Experience*, Harvard University Press, 1974.
- [17] M. Szell, R. Lambiotte, and S. Thurner, "Multirelational organization of large-scale social networks in an online world," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 31, pp. 13636–13641, 2010.
- [18] P. Klimek and S. Thurner, "Triadic closure dynamics drives scaling laws in social multiplex networks," *New Journal of Physics*, vol. 15, Article ID 063008, 2013.
- [19] B. Corominas-Murtra, B. Fuchs, and S. Thurner, "Detection of the elite structure in a virtual multiplex social system by means of a generalised K-core," *PLoS ONE*, vol. 9, no. 12, Article ID e112606, 2014.
- [20] L. Tang, X. Wang, and H. Liu, "Community detection via heterogeneous interaction analysis," *Data Min. Knowl. Discov.*, vol. 25, pp. 1–33, 2012.
- [21] T. G. Kolda, B. W. Bader, and J. P. Kenny, "Higher-order web link analysis using multilinear algebra," in *Proceedings of the 5th IEEE International Conference on Data Mining, ICDM 2005*, pp. 242–249, November 2005.
- [22] G. A. Barnett, H. W. Park, K. Jiang, C. Tang, and I. F. Aguillo, "A multi-level network analysis of web-citations among the worlds universities," *Scientometrics*, vol. 99, p. 26, 2014.
- [23] Z. Wu, W. Yin, J. Cao, G. Xu, and A. Cuzzocrea, "Community detection in multirelational social networks," in *Web Information Systems Engineering WISE 2013*, S. B. Heidelberg, Ed., vol. 8181 of *Lecture Notes in Computer Science*, pp. 43–56, 2014.
- [24] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, "Catastrophic cascade of failures in interdependent networks," *Nature*, vol. 464, no. 7291, pp. 1025–1028, 2010.
- [25] C. D. Brummitt, R. M. DSouza, and E. A. Leicht, "Suppressing cascades of load in interdependent networks," *Proc. Natl. Acad. Sci. USA*, vol. 109, pp. E680–E689, 2012.
- [26] A. Cardillo, M. Zanin, J. Gómez-Gardeñes, M. Romance, A. García del Amo, and S. Boccaletti, "Modeling the multi-layer nature of the european air transport network: Resilience and passengers re-scheduling under random failures," *Eur. Phys. J. Spec. Top.*, vol. 215, pp. 23–33, 2013.
- [27] M. Barigozzi, G. Fagiolo, and G. Mangioni, "Identifying the community structure of the international-trade multi-network," *Physica A: Statistical Mechanics and its Applications*, vol. 390, no. 11, pp. 2051–2066, 2011.
- [28] M. Montagna and C. Kok, *Multi-layered interbank model for assessing systemic risk*, Kiel, 2013.
- [29] R. Burkholz, M. V. Leduc, A. Garas, and F. Schweitzer, "Systemic risk in multiplex networks with asymmetric coupling and threshold feedback," *Physica D: Nonlinear Phenomena*, vol. 323–324, pp. 64–72, 2016.
- [30] D. J. Fenn, M. A. Porter, P. J. Mucha et al., "Dynamical clustering of exchange rates," *Quantitative Finance*, vol. 12, no. 10, pp. 1493–1520, 2012.
- [31] T. Di Matteo, F. Pozzi, and T. Aste, "The use of dynamical networks to detect the hierarchical organization of financial market sectors," *Eur. Phys. J. B*, vol. 73, pp. 3–11, 2010.
- [32] T. Aste, W. Shaw, and T. Di Matteo, "Correlation structure and dynamics in volatile markets," *New J. Phys.*, vol. 12, Article ID 085009, 2010.
- [33] C. Borghesi, M. Marsili, S. Miccichè, and S. Miccichè, "Emergence of time-horizon invariant correlation structure in financial returns by subtraction of the market mode," *Phys. Rev. E*, vol. 76, Article ID 026104, 2007.
- [34] N. Musmeci, T. Aste, and T. Di Matteo, "Relation between financial market structure and the real economy," *The PLOS ONE*, 2015.
- [35] J.-P. Onnela, A. Chakraborti, K. Kaski, and J. Kertész, "Dynamic asset trees and Black Monday," *Physica A: Statistical Mechanics and its Applications*, vol. 324, no. 1–2, pp. 247–252, 2003.
- [36] N. Musmeci, T. Aste, and T. Di Matteo, "Risk diversification: a study of persistence with a filtered correlation-network approach," *Journal of Network Theory in Finance*, vol. 1, pp. 1–22, 2015.
- [37] N. Musmeci, T. Aste, and T. Di Matteo, "What does past correlation structure tell us about the future? an answer from network theory," *Portfolio Management*, 2016.
- [38] V. Tola, F. Lillo, M. Gallegati, and R. N. Mantegna, "Cluster analysis for portfolio optimization," *Journal of Economic Dynamics and Control*, vol. 32, no. 1, pp. 235–258, 2008.
- [39] F. Pozzi, T. Di Matteo, and T. Aste, "Spread of risk across financial markets: better to invest in the peripheries," *Scientific Reports*, vol. 3, 2013.
- [40] H. Kaya, "Eccentricity in asset management," *The Journal of Network Theory in Finance*, vol. 1, no. 1, pp. 45–76, 2015.
- [41] D. Y. Kenett, M. Tumminello, A. Madi, G. Gur-Gershgoren, R. N. Mantegna, and E. Ben-Jacob, "Dominating clasp of the financial sector revealed by partial correlation analysis of the stock market," *PLoS ONE*, vol. 5, no. 12, Article ID e15032, 2010.
- [42] D. B. West, *Introduction to graph theory* (Prentice-Hall, Prentice-Hall, Englewood Cliffs, NJ, USA, 1996.
- [43] G. Bianconi, "Statistical mechanics of multiplex networks: entropy and overlap," *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, vol. 87, no. 6, Article ID 062806, 2013.
- [44] K. Pearson, "Note on regression and inheritance in the case of two parents," *Proceedings of the Royal Society of London*, vol. 58, pp. 240–242, 1895.
- [45] F. Pozzi, T. Di Matteo, and T. Aste, "Exponential smoothing weighted correlations," *Eur. Phys. J. B*, vol. 85, 2012.
- [46] R. Schmidt and U. Stadtmüller, "Nonparametric estimation of tail dependence," *Scandinavian Journal of Statistics*, vol. 33, pp. 307–335, 2006.
- [47] K. Baba, R. Shibata, and M. Sibuya, "Partial correlation and conditional correlation as measures of conditional independence," *Australian and New Zealand Journal of Statistics*, vol. 46, no. 4, pp. 657–664, 2004.

